Pathway-Based Genomics Prediction using Generalized Elastic Net.

Journal: PLoS Comput Biol

Publication Year: 2016

Authors: Artem Sokolov, Daniel E Carlin, Evan O Paull, Robert Baertsch, Joshua M Stuart

PubMed link: 26960204

Funding Grants: Center of Excellence for Stem Cell Genomics - Stanford

Public Summary:
Genetic markers have become a mainstay in modern cancer diagnosis. They enable a quickidentification of a tumor's subtype to suggest treatment options. Unfortunately, genetic markers provide little insight into a tumor's inner workings. Genes don't function in isolation, but rather belong to interacting parts, the collection of which define pathways. Knowing which pathways become dysregulated in cancer can help guide the development of drugs that shut down those pathways and kill tumor cells. However, gene markers are often selected in a way to produce the strongest diagnostic signal, with little consideration of how much the genes interact. This calls for new markers that leverage pathway structure. We introduced a method that selects gene markers in a way that maximizes the proximity of selected genes according to a provided gene-gene interaction network. As an analogy, if genes were people in a social network, our proposed method would build a team of diagnostic experts in a way that maximizes the number of friendships among the team members. Using experiments on synthetic data, we demonstrate that guiding marker selection towards genes that interact with each other maintains, and often improves, the accuracy of predictions even in cases where the full interaction network is unknown. We applied our method to predict the response of breast cancer cells to various chemical compounds and identified a novel pathway that may be related to cancer cells' resistance to a drug that targets a family of epidermal growth factor receptors.

Scientific Abstract:
We present a novel regularization scheme called The Generalized Elastic Net (GELnet) that incorporates gene pathway information into feature selection. The proposed formulation is applicable to a wide variety of problems in which the interpretation of predictive features using known molecular interactions is desired. The method naturally steers solutions toward sets of mechanistically interlinked genes. Using experiments on synthetic data, we demonstrate that pathway-guided results maintain, and often improve, the accuracy of predictors even in cases where the full gene network is unknown. We apply the method to predict the drug response of breast cancer cell lines. GELnet is able to reveal genetic determinants of sensitivity and resistance for several compounds. In particular, for an EGFR/HER2 inhibitor, it finds a possible trans-differentiation resistance mechanism missed by the corresponding pathway agnostic approach.